

Západočeská univerzita v Plzni

Fakulta aplikovaných věd

Katedra matematiky

BAKALÁŘSKÁ PRÁCE

Statistická analýza dat týkajících se
nezaměstnanosti

Plzeň, 2015

Lenka Malínská

Prohlášení

Prohlašuji, že jsem svou bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů uvedených v příloženém seznamu.

V Plzni dne podpis

Poděkování

Ráda bych poděkovala vedoucímu mé bakalářské práce, doc. RNDr. Petru Stehlíkovi, Ph.D., za jeho odborné rady, poskytnutou literaturu a hlavně za čas, který mi při přípravě práce věnoval. Dále děkuji svojí rodině za trpělivost a pomoc, kterou mi po dobu celého studia poskytovali.

Abstrakt

Cílem této práce je statistické zpracování dat z oblasti automotive. Data shromážděná ze všech států Evropské unie mezi roky 2000 až 2014 v první části popisujeme, dále provádíme základní analýzu dat, na základě které stanovíme několik zajímavých hypotéz, které v poslední kapitole pomocí statistických metod ověřujeme. V rámci testování hypotéz zjišťujeme vliv ekonomické krize na zaměstnanost a produkci v oblasti automobilového průmyslu. Také prověřujeme hypotézy o přímém dopadu výše automobilové produkce na ekonomiku jako celku v České republice i v jiných zemích Evropské unie.

Klíčová slova: automotive, zaměstnanost, lineární regrese, klamná korelace, intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování, předpoklady lineární regrese

Abstract

The aim of this thesis is statistical data processing in the field of automotive. In the first part we describe the data which has been collected from all European Union countries between the years 2000-2014. Based on providing of basic data analysis, we set several interesting hypotheses that has been verified in the last chapter using statistical methods. Within hypothesis testing, we find the influence of the economic crisis on employment and production in the automotive industry. We also verify the hypothesis of a direct impact on amount of automobile production on the global economics in the Czech Republic and other European Union countries.

Keywords: automotive, employment, linear regression, illusory correlation, regression confidence intervals for values and for future observations, assumptions of linear regression

Obsah

1	Úvod	1
2	Popis dat.....	2
2.1	Datový soubor.....	2
2.2	Struktura dat.....	2
3	Základní zpracování dat.....	6
4	Testování dat.....	9
4.1	Lineární regrese HDP/obyvatele na produkci aut	9
4.2	Klamná korelace	10
4.3	Lineární regrese HDP/obyvatele na produkci aut v souvislosti s ekonomickou krizí...11	
4.4	Lineární regrese zaměstnanosti v automobilovém průmyslu na zaměstnanosti mimo automobilový průmysl	14
4.5	Intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování	16
4.6	Lineární regrese produkce osobních a nákladních automobilů v EU a ověřování předpokladů regrese	17
4.6.1	Předpoklad nulovosti střední hodnoty náhodných odchylek	19
4.6.2	Předpoklad konstantního rozptylu náhodných odchylek	19
4.6.3	Předpoklad nezávislosti náhodných odchylek	20
5	Závěr.....	22
6	Citovaná literatura.....	24
7	Obsah příloženého CD.....	25

1 Úvod

V této práci jsme se zaměřili na statistickou analýzu dat týkajících se oblasti automotive, tedy automobilového průmyslu, který v posledních desetiletích zaznamenal výrazný nárůst produkce a zaměstnanosti nejen celosvětově, v Evropské unii, ale i v České republice, a je tedy velmi zajímavou oblastí pro analýzu.

V současné době automobilový průmysl v České republice představuje více než pětinu průmyslové výroby a zaměstnává více než 100 tisíc pracovníků. Je tedy klíčovým oborem pro národní hospodářství. Rozhodli jsme se ověřit, zda automotive, považovaný také v Německu za motor ekonomiky, přímo ovlivňuje ekonomiku evropských států a zda citlivě reaguje na změny v ekonomice, například na ekonomickou krizi. (Sůra, 2011)

Pro zpracování byla shromážděna data z posledních dvou desetiletí, kdy byl vývoj automobilového průmyslu nejrychlejší a z tohoto období bylo možno zároveň získat větší množství dat, vhodných pro statistickou analýzu. Datový soubor detailně popíšeme v první části této práce.

V následující kapitole se zabýváme základním statistickým zpracováním dat, na jehož základě stanovíme zajímavé hypotézy, které budeme v další kapitole ověřovat.

Ve 4. kapitole tedy formulujeme a ověřujeme statistické hypotézy pomocí lineární regrese. Při ověřování hypotéz jsme zároveň využili různých statistických metod, které v této kapitole blíže ukážeme. U první hypotézy jsme narazili na zajímavý jev, tzv. klamnou korelaci, kterou jsme se rozhodli testovat. K hypotéze, ve které se věnujeme ověřování korelace zaměstnanosti v automotive a mimo něj, zkonstruujeme intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování a v posledním odstavci této kapitoly se zabýváme ověřováním předpokladů lineární regrese.

V závěru práce popíšeme výsledky a shrneme zajímavé poznatky, které jsme v průběhu zpracování objevili.

2 Popis dat

2.1 Datový soubor

Datový soubor tvoří data získaná z databáze evropského statistického úřadu Eurostat (Statistický úřad Evropské unie, 2015).

Eurostat je statistický úřad Evropských společenství se sídlem v Lucembursku založený v roce 1953. Od roku 1958 je Generálním ředitelstvím (DG – Directorate-General) Evropské komise. Klíčovým úkolem Eurostatu je poskytování statistik jiným generálním ředitelstvím, poskytování dat Komisi a ostatním evropským institucím za účelem definování, realizace a analýzy politiky Společenství.

Úkolem Eurostatu je poskytovat statistiky na evropské úrovni umožňující srovnání mezi jednotlivými zeměmi a regiony. Mezinárodní statistiky informují o stavu a vývoji společnosti nejen v členských zemích EU, ale i v dalších státech. Jsou důležitým, objektivním a realistickým způsobem měření toho, jak žijeme. Databáze Eurostatu nabízejí široké veřejnosti zdarma řadu významných a zajímavých statistických údajů. Mezi nejčastější uživatele patří např. mezinárodní i vládní instituce, podniky, novináři a studenti. (Český statistický úřad, 2015)

Existují dvě hlavní rozhraní pro přístup a extrakci dat z databáze Eurostat:

1. Databáze podle témat (Datový navigační strom) - údaje, které jsou přístupné prostřednictvím nástrojů pro extrakci, kde odborník může najít sofistikované informace a definovat vlastní parametry.
2. Tabulky podle témat – předdefinované, již hotové tabulky, obsahující základní údaje o každém tématu, jakož i údaje o klíčových ukazatelích politiky EU. Zde mohou i příležitostní uživatelé najít údaje, které hledají.

Pro účely této práce byla data vyhledávána a extrahována podle rozhraní ad 1. ve formátu .xls a následně zpracovávána v programu MS Excel.

2.2 Struktura dat

Jednotným parametrem pro vyhledávání dat byla geografická poloha.

Pro účely této práce byla vyhledávána data ze všech 28 členských států Evropské unie (dále jen EU) a také pro EU jako celek.

Data jsme vyhledávali v časových obdobích a časových intervalech. Oba parametry jsou závislé na dostupnosti dat v databázi, proto jsou data v těchto parametrech odlišná.

Časové období jsme tedy přizpůsobovali a volili vždy stejně tak, aby bylo možno jednotlivá data společně analyzovat. Např. data o nezaměstnanosti jsou v databázi pro jednotlivé členské státy EU dostupná již od r. 1998 do r. 2015 oproti tomu data o zaměstnanosti v automobilovém průmyslu jsou dostupná od r. 2002 do r. 2012.

Časový interval byl vždy volen podle dostupnosti dat nejkratší možný, aby bylo dosaženo maximální velikosti statistického souboru. Dostupné intervaly pro námi analyzovaná data byly roční případně čtvrtletní.

Data byla shromažďována z následujících ekonomických oblastí:

- Hrubý domácí produkt (dále jen HDP), uváděný v mil. EUR
- Počet obyvatel
- HDP/obyvatele, uváděný v EUR/obyvatele
- Zaměstnanost celkem, kde zaměstnanost je definována jako počet osob starších 15-ti let, které během referenčního týdne odpracovaly alespoň jednu hodinu za mzdu nebo zisk, a také osob, které byly v zaměstnání dočasně nepřítomné. (Statistický úřad Evropské unie, 2015)
- Zaměstnanost v automobilovém průmyslu
- Produkce osobních a nákladních automobilů
- Index produkce osobních a nákladních automobilů, kde index 100 je roven průměru hodnot za všechna čtvrtletí roku 2010

Roční data jsme získali pro všechny země EU pro tyto oblasti a období:

- Hrubý domácí produkt (dále jen HDP), od r. 2002 do r. 2013
- Počet obyvatel, od r. 2002 do r. 2013
- Zaměstnanost celkem, od r. 2002 do r. 2012
- Zaměstnanost v automobilovém průmyslu, od r. 2002 do r. 2012
- Produkce osobních a nákladních automobilů, od r. 2000 do r. 2012

Kvartální data se nám podařilo shromáždit pro všechny země EU pro níže uvedené oblasti a období:

- Index produkce osobních a nákladních automobilů, od 1. čtvrtletí roku 2003 do 3. čtvrtletí roku 2014
- Index produkce automobilů celkem, od 1. čtvrtletí roku 2001 do 1. čtvrtletí roku 2014
- HDP/obyvatele, od 1. čtvrtletí roku 2001 do 1. čtvrtletí roku 2014

Pro zpracování dat z automobilového průmyslu bylo nejprve třeba definovat jednotný parametr pro vyhledávání a následnou extrakci dat.

Podle statistické klasifikace ekonomických aktivit Evropského společenství byl jako jednotný parametr zvolen kód NACE 29.10, tedy výroba motorových vozidel.

Statistical Classification of Economic Activities in the European Community, Rev. 2 (2008)

-> C: MANUFACTURING	
-> 29: Manufacture of motor vehicles, trailers and semi-trailers	
-> 29.1: Manufacture of motor vehicles	
CODE:	29.10
DESCRIPTION:	Manufacture of motor vehicles
REFERENCE TO ISIC REV. 4:	2910
THIS ITEM INCLUDES:	<p>This class includes:</p> <ul style="list-style-type: none">- manufacture of passenger cars- manufacture of commercial vehicles:<ul style="list-style-type: none">• vans, lorries, over-the-road tractors for semi-trailers etc.- manufacture of buses, trolley-buses and coaches- manufacture of motor vehicle engines- manufacture of chassis for motor vehicles- manufacture of other motor vehicles:<ul style="list-style-type: none">• snowmobiles, golf carts, amphibious vehicles• fire engines, street sweepers, travelling libraries, armoured cars etc.• concrete-mixer lorries- ATVs, go-carts and similar including race cars
THIS ITEM ALSO INCLUDES:	<p>This class also includes:</p> <ul style="list-style-type: none">- factory rebuilding of motor vehicle engines
THIS ITEM EXCLUDES:	<p>This class excludes:</p> <ul style="list-style-type: none">- manufacture of electric motors (except starting motors), see 27.11- manufacture of lighting equipment for motor vehicles, see 27.40- manufacture of pistons, piston rings and carburettors, see 28.11- manufacture of agricultural tractors, see 28.30- manufacture of tractors used in construction or mining, see 28.92- manufacture of off-road dumping trucks, see 28.92- manufacture of bodies for motor vehicles, see 29.20- manufacture of electrical parts for motor vehicles, see 29.31- manufacture of parts and accessories for motor vehicles, see 29.32- manufacture of tanks and other military fighting vehicles, see 30.40- maintenance and repair of motor vehicles, see 45.20

Obrázek 2.1 Detailní statistická klasifikace kódu NACE 29.10

K získání detailních statistik podle klasifikace viz výše, tedy až do čtyřmístného kódu NACE, byla kontaktována centrální podpora Eurostatu. Poděkování patří paní Katarzyně Kraszewské z oddělení ESDS (Podpora evropských statistických dat), která nám poskytla informace k vyhledání požadovaných statistik.

Unemployment by sex and age groups - quarterly average, 1 000 persons [une_nb_q]

Last update 22.10.13
 Extracted on 28.10.13
 Source of data Eurostat
 Short Description Short Description is not available
 S_ADJ Seasonally adjusted data
 AGE Total
 SEX Total

GEO/TIME	2003Q1	2003Q2	2003Q3	2003Q4	2004Q1	2004Q2
European Union (28 countries)	20 800	20 919	20 919	21 030	21 389	21 342
Belgium	356	360	357	374	382	358
Bulgaria	503	469	438	415	420	412
Czech Republic	378	396	408	418	432	428
Denmark	143	158	156	166	162	157
Germany	3 832	3 945	3 947	3 936	3 990	4 140
Estonia	69	72	67	64	66	62
Ireland	86	85	90	85	94	86
Greece	458	458	458	469	521	505
Spain	2 244	2 226	2 258	2 247	2 230	2 250
France	2 410	2 466	2 436	2 547	2 603	2 554
Croatia	251	254	253	252	253	253
Italy	2 108	2 064	2 036	1 993	1 987	1 953
Cyprus	13	14	14	15	15	16
Latvia	111	111	120	107	118	108
Lithuania	198	211	200	189	189	181
Luxembourg	6	7	8	8	9	10
Hungary	236	239	238	237	240	247
Malta	12	12	13	13	12	11
Netherlands	299	331	353	378	400	431
Austria	159	163	170	172	195	190
Poland	3 330	3 303	3 319	3 305	3 352	3 239
Portugal	375	387	383	391	376	402
Romania	678	680	669	720	814	809
Slovenia	63	65	65	64	64	63
Slovakia	463	450	454	465	490	491
Finland	238	238	234	232	232	232
Sweden	291	295	307	329	343	345
United Kingdom	1 491	1 462	1 469	1 437	1 398	1 407

Special value:

: not available

Obrázek 2.2 Ukázka datového souboru – Nezaměstnanost v zemích EU

3 Základní zpracování dat

V této části se budeme věnovat základní analýze získaných dat. Na základě výsledků této analýzy v následující kapitole stanovíme zajímavé hypotézy, které budeme testovat.

Nejprve jsme určili základní statistické charakteristiky.

Pro jednotlivé země EU a sledované roční ekonomické ukazatele jsme podrobněji určili průměry a mediány. Tyto hodnoty jsou uvedeny v tabulkách 3.1 Hodnoty průměru ročních dat zemí EU v letech 2002 až 2013 a 3.2 Hodnoty mediánu ročních dat zemí EU v letech 2002 až 2013.

státy	počet obyvatel	HDP	HDP/obyvatele	míra nezaměstnanosti	zaměstnanost	zaměstnanost v automotive	podíl zaměstnanosti v automotive na zaměstnanosti
Belgie	10 676 728	330 518	30 890	7,9%	4 287 536	53 382	1,2%
Bulharsko	7 581 565	30 159	4 010	10,6%	3 006 573	5 073	0,2%
Česká republika	10 333 786	126 592	12 223	6,9%	4 772 582	136 745	2,9%
Dánsko	5 475 100	221 136	40 347	5,5%	2 704 173	5 073	0,2%
Německo	82 200 065	2 399 913	29 203	8,5%	37 118 745	1 061 327	2,9%
Estonsko	1 346 024	13 627	10 148	9,8%	590 264	2 509	0,4%
Irsko	4 318 823	162 028	37 521	8,1%	1 880 273	4 100	0,2%
Řecko	11 100 010	200 891	18 086	11,8%	4 231 600	2 427	0,1%
Španělsko	44 673 840	964 763	21 531	14,2%	18 385 418	218 536	1,2%
Francie	63 683 336	1 837 103	28 805	9,1%	25 087 309	310 818	1,2%
Chorvatsko	4 300 506	39 970	9 296	12,3%	1 503 491	2 855	0,2%
Itálie	59 021 583	1 489 836	25 223	8,0%	22 372 727	212 355	0,9%
Kypr	778 099	15 266	19 527	5,6%	353 982	0	0,0%
Litva	2 183 475	17 179	7 965	11,9%	971 909	0	0,0%
Lotyšsko	3 228 889	25 762	8 089	10,9%	1 387 455	0	0,0%
Lucembursko	483 909	35 091	71 984	4,5%	203 973	0	0,0%
Maďarsko	10 049 283	90 947	9 057	8,2%	3 842 682	66 545	1,7%
Malta	407 648	5 717	13 990	6,8%	155 636	0	0,0%
Nizozemí	16 432 126	551 213	33 516	4,2%	8 205 209	20 436	0,2%
Rakousko	8 279 875	268 626	32 396	4,5%	3 893 727	34 755	0,9%
Polsko	38 260 680	300 182	7 842	13,2%	14 595 491	157 791	1,1%
Portugalsko	10 511 983	161 062	15 317	9,7%	4 733 936	42 745	0,9%
Rumunsko	20 872 913	104 289	5 044	7,1%	8 818 400	97 318	1,1%
Slovinsko	2 020 911	32 216	15 924	6,5%	928 427	13 573	1,5%
Slovensko	5 383 414	52 769	9 797	14,5%	2 269 782	56 691	2,5%
Finsko	5 297 626	171 508	32 337	8,0%	2 416 327	7 536	0,3%
Švédsko	9 186 022	331 875	36 035	7,3%	4 378 591	66 455	1,5%
Velká Británie	61 397 031	1 820 539	29 662	6,1%	28 089 082	231 827	0,8%
EU celkem	499 485 251	11 800 777	23 606	8,6%	211 185 300	2 810 873	1,3%

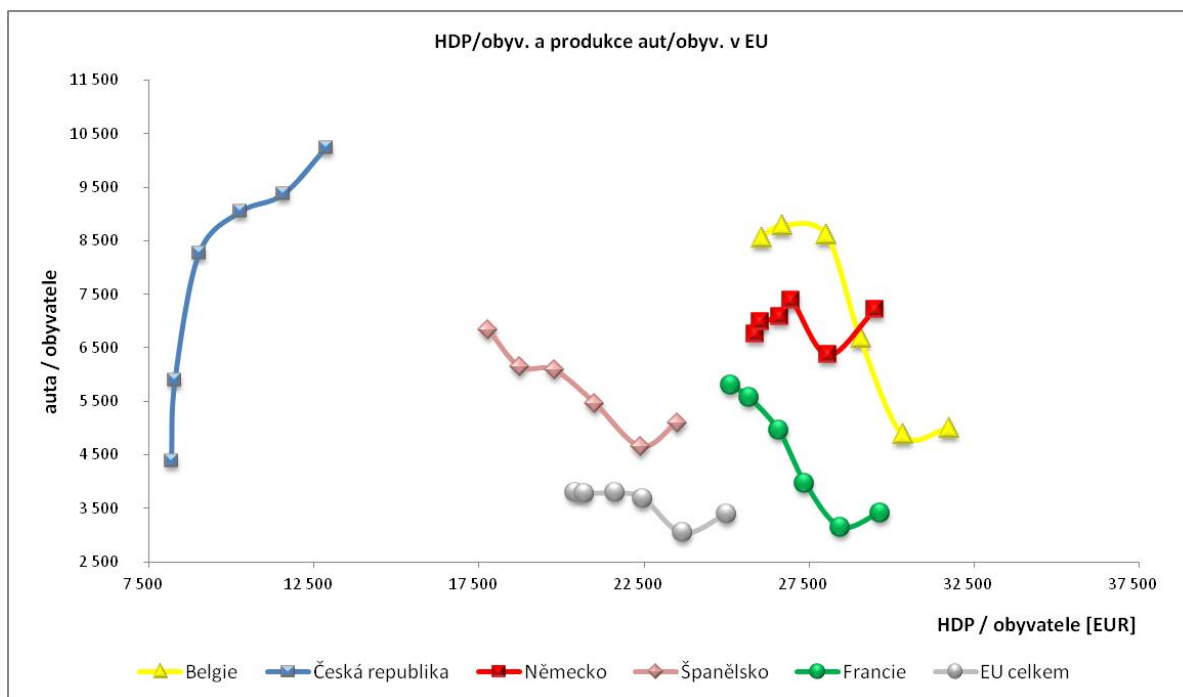
Tabulka 3.1 Hodnoty průměru ročních dat zemí EU v letech 2002 až 2013

státy	počet obyvatel	HDP	HDP/obyvatele	míra nezaměstnanosti	zaměstnanost	zaměstnanost v automotive	podíl zaměstnanosti v automotive na zaměstnanosti
Belgie	10 625 700	338 242	31 704	7,9%	4 348 100	52 600	1,2%
Bulharsko	7 545 338	32 853	4 371	10,3%	2 947 000	4 300	0,1%
Česká republika	10 298 828	137 053	13 251	7,1%	4 796 400	143 100	3,0%
Dánsko	5 461 438	225 555	41 169	5,4%	2 705 300	5 300	0,2%
Německo	82 266 372	2 401 350	29 228	8,7%	37 397 200	1 072 700	2,9%
Estonsko	1 342 330	14 170	10 587	10,1%	586 300	2 900	0,5%
Irsko	4 398 942	162 748	35 988	4,7%	1 837 500	4 000	0,2%
Řecko	11 117 574	201 140	18 090	9,9%	4 286 700	2 600	0,1%
Španělsko	45 226 803	1 024 444	22 186	11,4%	18 304 100	221 800	1,2%
Francie	63 826 129	1 886 278	29 475	9,3%	25 425 100	323 500	1,3%
Chorvatsko	4 305 610	43 642	10 162	12,8%	1 511 500	3 400	0,2%
Itálie	59 262 747	1 535 790	25 566	8,0%	22 496 500	218 800	1,0%
Kypr	767 125	16 157	20 234	4,6%	367 900	0	0,0%
Litva	2 200 325	18 280	8 535	10,7%	979 700	0	0,0%
Lotyšsko	3 231 294	27 182	8 595	12,4%	1 416 200	0	0,0%
Lucembursko	479 993	36 474	74 772	4,6%	201 800	0	0,0%
Maďarsko	10 055 780	93 829	9 362	7,5%	3 849 200	61 700	1,6%
Malta	406 724	5 766	14 120	6,9%	155 500	0	0,0%
Nizozemí	16 381 696	572 504	34 863	4,4%	8 226 900	21 200	0,3%
Rakousko	8 300 788	275 124	33 071	4,4%	3 963 200	34 700	0,9%
Polsko	38 182 222	310 842	8 152	10,1%	14 996 500	187 500	1,3%
Portugalsko	10 522 288	165 224	15 714	8,6%	4 800 000	36 400	0,8%
Rumunsko	20 882 982	121 262	5 843	7,2%	8 822 000	83 000	0,9%
Slovinsko	2 010 323	34 774	17 080	6,3%	924 900	15 800	1,7%
Slovensko	5 377 508	58 803	10 934	14,0%	2 303 200	53 000	2,3%
Finsko	5 288 720	175 521	32 875	8,2%	2 416 100	7 400	0,3%
Švédsko	9 148 092	325 713	35 728	7,4%	4 391 400	66 000	1,5%
Velká Británie	61 296 895	1 811 712	29 687	5,4%	28 183 500	254 700	0,9%
EU celkem	500 331 796	12 076 389	24 055	8,8%	213 016 800	2 840 100	1,3%

Tabulka 3.2 Hodnoty mediánu ročních dat zemí EU v letech 2002 až 2013

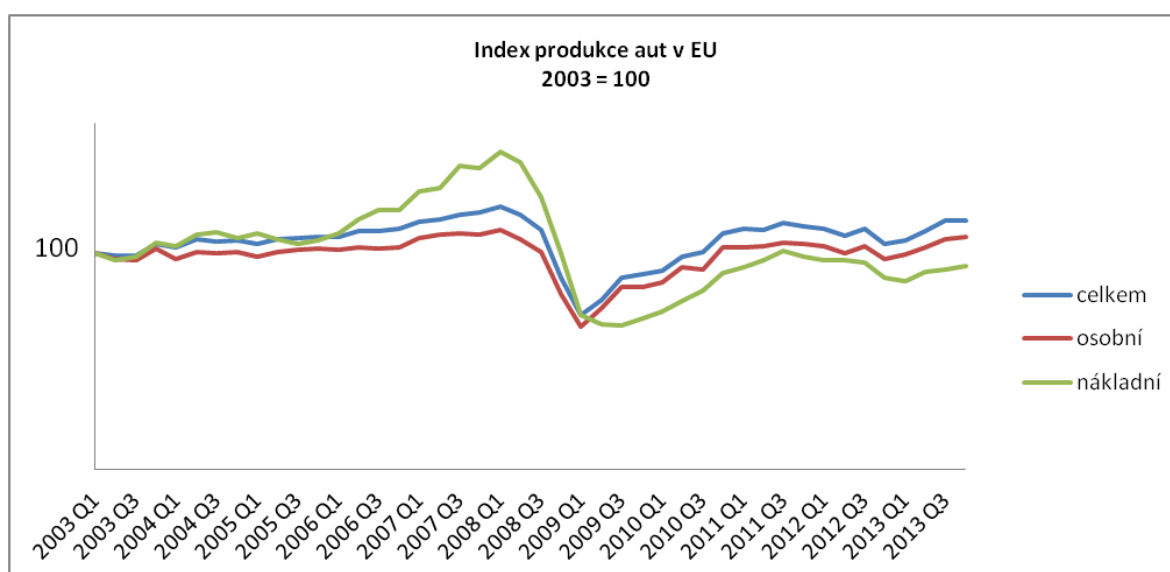
Z malých rozdílů mezi hodnotami průměrů a mediánů lze usoudit, že data jsou poměrně symetricky rozložena okolo průměru a neobsahují odlehlá pozorování. V porovnání se zeměmi EU jako celku data v zemích EU jednotlivě nevykazují tak vysokou variabilitu.

Z dat v letech 2006-2012 o počtu obyvatel v EU a výši HDP byly vypočteny hodnoty HDP na obyvatele, stejným způsobem jsme získali produkci aut přepočtenou na obyvatele. Grafické zpracování výsledků na vybraných členských státech EU a také EU celkem je zobrazeno na obrázku 3.1 Zobrazení závislosti HDP/obyv. na produkci aut/obyv. v Belgii, ČR, Německu, Španělsku, Francii a EU celkem v letech 2006 až 2012. Pro ilustraci časové posloupnosti jsme jednotlivé body v grafu spojili. U České republiky je možno pozorovat růst obou veličin v čase. Tento trend budeme v další kapitole hlouběji analyzovat.



Obrázek 3.1 Zobrazení závislosti HDP/obyv. na produkci aut/obyv. v Belgii, ČR, Německu, Španělsku, Francii a EU celkem v letech 2006 až 2012

Vývoj produkce osobních a nákladních automobilů v jednotlivých čtvrtletích let 2003-2014 je zobrazen na obrázku 3.3 Index produkce osobních a nákladních aut v EU v jednotlivých čtvrtletích let 2003-2014. V grafu můžeme sledovat velký nárůst produkce nákladních automobilů v době před ekonomickou krizí, oproti poměrně stabilní produkci osobních aut. V 1. čtvrtletí r. 2009 dochází k velkému propadu produkce aut nákladních, resp. osobních. V dalších letech se produkce opět stabilizuje a poměrně konstantně roste u obou sledovaných proměnných. Propad produkce by mohl být odezvou na ekonomickou krizi, která začala v r. 2008 a mohla výrazně ovlivňovat produkci automobilového průmyslu na začátku dalšího roku.



Obrázek 3.2 Index produkce osobních a nákladních aut v EU v jednotlivých čtvrtletích let 2003-2014

4 Testování dat

Z výsledků základní analýzy vyplynuly zajímavé domněnky, které budeme v této části hlouběji analyzovat.

Pro výpočty lineární regrese bylo použito nástroje Analýza dat v programu MS Excel.

4.1 Lineární regrese HDP/obyvatele na produkci aut

Pro testování závislosti výše reálného HDP/obyvatele v České republice na produkci v automobilovém průmyslu jsme zvolili lineární regresi.

Uvažujeme následující regresní model:

$$y_i = \theta_0 + \theta_1 x_i + \varepsilon_i \quad (i = 1, 2, \dots, n),$$

kde parametry θ_0 a θ_1 jsou neznámé parametry, které budeme odhadovat a ε_i je neznámá náhodná odchylka. Parametry θ_j tohoto modelu jsou zastoupeny v lineárním tvaru. (Reif, 2004)

V našem případě je vysvětlující proměnnou (x_i) HDP České republiky, přepočtený na obyvatele a vysvětlovanou proměnnou (y_i), jejíž parametry odhadujeme, je objemový index produkce automobilů v České republice. Obě veličiny porovnáváme za stejné časové období, od 1. čtvrtletí r. 2000 do 1. čtvrtletí r. 2014.

Testujeme hypotézu:

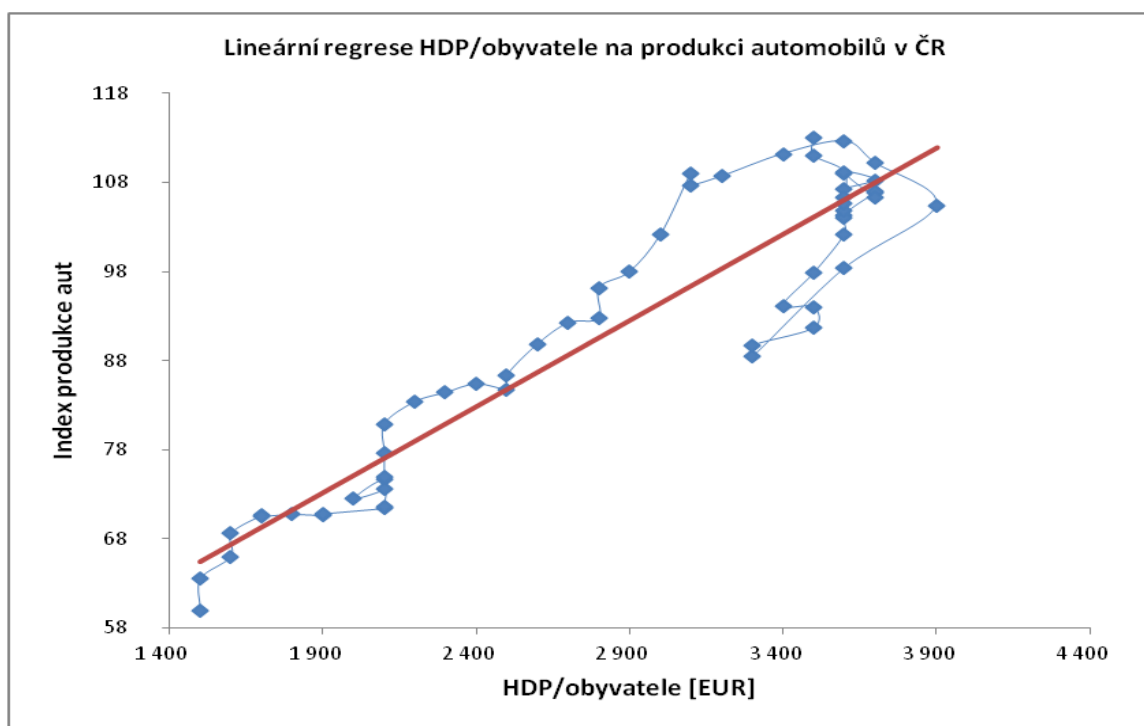
$$H_0 : \theta_1 = 0$$

Výsledek analýzy je uveden v tabulce 4.1 Výsledek lineární regrese HDP/obyv. na produkci v automobilovém průmyslu v ČR v letech 2000-2014.

Hodnota F	381,05227
Významnost F	$2,14108 \cdot 10^{-26}$
Hladina významnosti α	0,05
Koeficient β_1	0,01935
Korelační koeficient	0,93481

Tabulka 4.1 Výsledek lineární regrese HDP/obyv. na produkci v automobilovém průmyslu v ČR v letech 2000-2014

Grafické znázornění výsledku analýzy je zobrazeno na obrázku 4.1 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2000-2014.



Obrázek 4.1 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2000-2014

Na základě výsledků analýzy zamítáme hypotézu H_0 na hladině významnosti $\alpha = 5\%$ a konstatujeme, že model jako celek je statisticky významný, protože významnost testu F je menší než hladina významnosti α .

Podle hodnoty korelačního koeficientu se jedná o silnou kladnou závislost.

Abychom vyloučili, že se nejedná o klamnou korelaci, tedy že produkce automobilů je přímo závislá na HDP/obyvatele, budeme data dále testovat.

4.2 Klamná korelace

V této části budeme blíže zkoumat jev klamné korelace (Reif, 2004), kdy hodnota korelačního koeficientu mezi produkcí automobilů v ČR a HDP/obyvatele v ČR může být vysoká nikoliv vlivem přímé souvislosti mezi těmito veličinami, ale proto, že veličiny jsou ovlivněny jinou náhodnou veličinou, v našem případě časem.

Testujeme významnost výběrového koeficientu parciální korelace $r_{XY.Z}$, tedy korelačního koeficientu mezi rezidui, která vzniknou při lineární regresní závislosti HDP/obyvatele v čase a produkci aut v čase, tzn. nezávislost veličin.

Pro výběrový koeficient parciální korelace $r_{XY.Z}$ platí, že:

$$r_{XY.Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{(1 - r_{XZ}^2)(1 - r_{YZ}^2)}}$$

Testujeme hypotézu:

$$H_0 : \rho = 0$$

proti oboustranné alternativě $\rho \neq 0$.

Testovací statistika má tvar:

$$t = \frac{r_{XY.Z}}{\sqrt{1 - r_{XY.Z}^2}} \sqrt{n - 2}$$

Hodnoty vypočtených korelačních koeficientů a výběrového koeficientu parciální korelace jsou shrnuty v tabulce 4.2 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem.

r_{xy}	0,935
r_{xz}	0,944
r_{yz}	0,892
$r_{xy.z}$	0,621

Tabulka 4.2 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem

Shrnutí výsledku testu klamné korelace je uveden v tabulce 4.3 Výsledek testu nezávislosti HDP/obyv. a produkci aut na čase.

$n-3$	α	$t_{1-\alpha/2}(n-3)$	$r_{xy.z}$	t
54	5%	2,00	0,62	5,83

Tabulka 4.3 Výsledek testu nezávislosti HDP/obyv. a produkci aut na čase

Na základě hodnoty testovací statistiky t , hypotézu H_0 zamítáme na hladině významnosti $\alpha = 5\%$ a konstatujeme, že výběrový koeficient parciální korelace je významný a HDP/obyvatele a produkce aut v ČR jsou i při zahrnutí vlivu času na sobě závislé.

V grafu na obrázku 4.1 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2000-2014, můžeme pozorovat růst obou veličin v čase, v určitém bodě se však křivka mění a lineární růst se zastavuje. Zjistili jsme, že bod, ve kterém se křivka začíná měnit odpovídá počátku ekonomické krize v roce 2008. Rozhodli jsme se tedy data rozdělit na dobu před ekonomickou krizí a v krizi a znovu je otestovat.

4.3 Lineární regrese HDP/obyvatele na produkci aut v souvislosti s ekonomickou krizí

Testujeme závislost výše reálného HDP/obyvatele v České republice na produkci v automobilovém průmyslu před ekonomickou krizí, tedy v letech 2000 až 2007, resp. v době této krize v letech 2008 až 2009 (Pavlínek & Ženka, 2010).

Výsledek testu lineární regrese před ekonomickou krizí je uveden v tabulce 4.4 Výsledek lineární regrese HDP/obyv. v ČR na produkci v automobilovém průmyslu v letech 2000-2007.

Hodnota F	709,33983
Významnost F	$1,96038 \cdot 10^{-22}$
Hladina významnosti α	0,05
Koeficient β_1	0,02654
Korelační koeficient	0,97951

Tabulka 4.4 Výsledek lineární regrese HDP/obyv. v ČR na produkci v automobilovém průmyslu v letech 2000-2007

Hypotézu $H_0 : \beta_1 = 0$ zamítáme na hladině významnosti $\alpha = 5\%$ a konstatujeme, že model jako celek je vysoce statisticky významný pro roky 2000 - 2007. Po provedení testu nezávislosti veličin zjišťujeme, že se opět nejedná o klamnou korelaci a veličiny jsou na sobě závislé s vyloučením vlivu času. V tabulce 4.5 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem v letech 2000-2007 a tabulce 4.6 Výsledek testu nezávislosti HDP/obyv. a produkci aut v letech 2000-2007 na čase nalezneme příslušné hodnoty.

r_{xy}	0,980
r_{xz}	0,982
r_{yz}	0,968
r_{xy-z}	0,610

Tabulka 4.5 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem v letech 2000-2007

n-3	α	$t_{1-\alpha/2}(n-3)$	r_{xy-z}	t
29	5%	2,05	0,61	4,00

Tabulka 4.6 Výsledek testu nezávislosti HDP/obyv. a produkci aut v letech 2000-2007 na čase

Pro roky 2008 až 2009, tedy dobu ekonomické krize, zamítáme hypotézu $H_0 : \beta_1 = 0$ na hladině významnosti $\alpha = 5\%$, produkce v automobilovém průmyslu v ČR tedy v době krize závisela na výši reálného HDP/obyvatele. Nicméně na základě provedeného testu významnosti výběrového koeficientu parciální korelace r_{xy-z} konstatujeme, že výběrový koeficient parciální korelace není významný.

Výsledky obou testů a hodnoty korelačních koeficientů jsou uvedeny v tabulce 4.7 Výsledek lineární regrese HDP/obyv. v ČR na produkci v automobilovém průmyslu v letech 2008-2009, tabulce 4.8 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem v letech 2008-2009a tabulce 4.9 Výsledek testu nezávislosti HDP/obyv. a produkci aut v letech 2008-2009 na čase.

Hodnota F	8,53126
Významnost F	0,02660
Hladina významnosti α	0,05
Koeficient β_1	0,03629
Korelační koeficient	0,76622

Tabulka 4.7 Výsledek lineární regrese HDP/obyv. v ČR na produkci v automobilovém průmyslu v letech 2008-2009

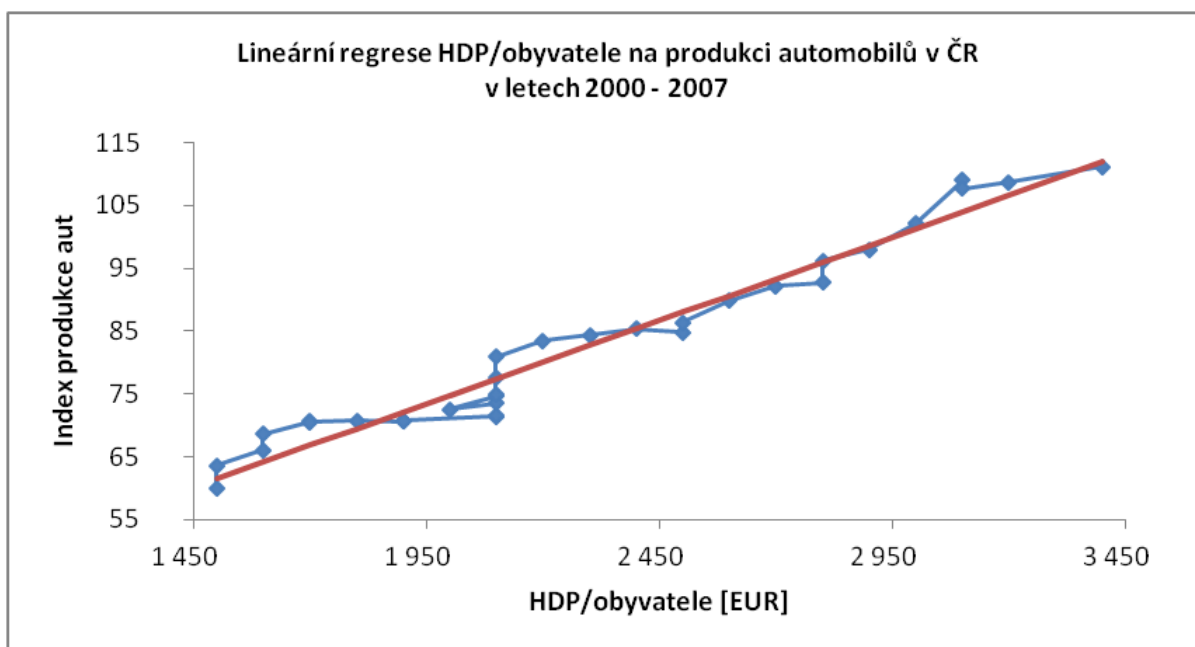
r_{XY}	0,766
r_{XZ}	-0,554
r_{YZ}	-0,861
$r_{XY \cdot Z}$	0,683

Tabulka 4.8 Hodnoty korelačních koeficientů mezi produkcí automobilů v ČR a HDP/obyvatele v ČR a časem v letech 2008-2009

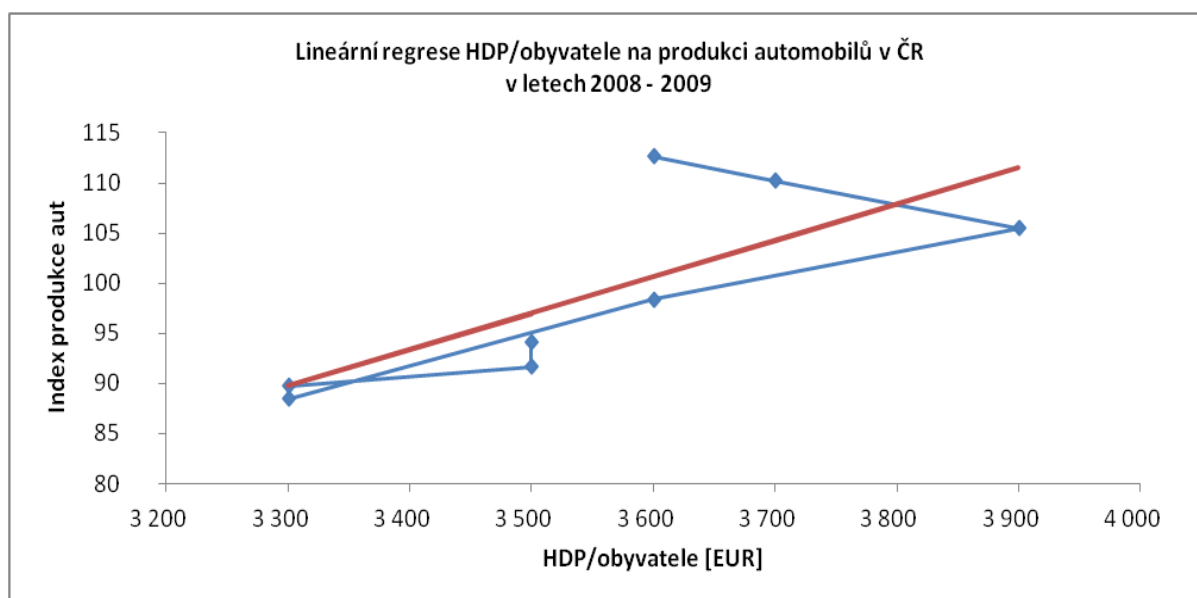
n-3	α	$t_{1-\alpha/2}(n-3)$	$r_{XY \cdot Z}$	t
5	5%	2,571	0,683	1,619

Tabulka 4.9 Výsledek testu nezávislosti HDP/obyv. a produkci aut v letech 2008-2009 na čase

Grafické zobrazení výsledků obou analýz nalezneme na obrázku 4.2 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2000-2007 resp. 4.3 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2008-2009.



Obrázek 4.2 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2000-2007



Obrázek 4.3 Lineární regrese HDP/obyvatele na produkci automobilů v ČR v letech 2008-2009

4.4 Lineární regrese zaměstnanosti v automobilovém průmyslu na zaměstnanosti mimo automobilový průmysl

V tomto odstavci chceme ověřit hypotézu korelace změn zaměstnanosti v automobilovém průmyslu a mimo něj pro země EU a pro Německo pomocí lineární regrese a také zkonstruovat intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování.

Pro lineární regresní model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ jsme opět stanovili hypotézu:

$$H_0 : \beta_1 = 0$$

V tomto případě vysvětlujeme změny zaměstnanosti mimo automobilový průmysl (x_i) změnami zaměstnanosti v automobilovém průmyslu (y_i). Vzhledem k tomu, že kvartální data dostupná na Eurostatu (Statistický úřad Evropské unie, 2015) byla neúplná, použili jsme dostupná roční data od r. 2003 do r. 2012. Malá velikost statistického souboru nicméně může mít vliv na výsledek našeho testování. Výsledek tohoto testování je uveden v tabulce 4.10 Výsledek lineární regrese změn zaměstnanosti mimo automobilový průmysl na změnách zaměstnanosti v automobilovém průmyslu v EU a 4.11 Výsledek lineární regrese změn zaměstnanosti mimo automobilový průmysl na změnách zaměstnanosti v automobilovém průmyslu v Německu.

Hodnota F	2,15353
Významnost F	0,18042
Hladina významnosti α	0,05
Koeficient β_1	0,09160
Korelační koeficient	0,46054

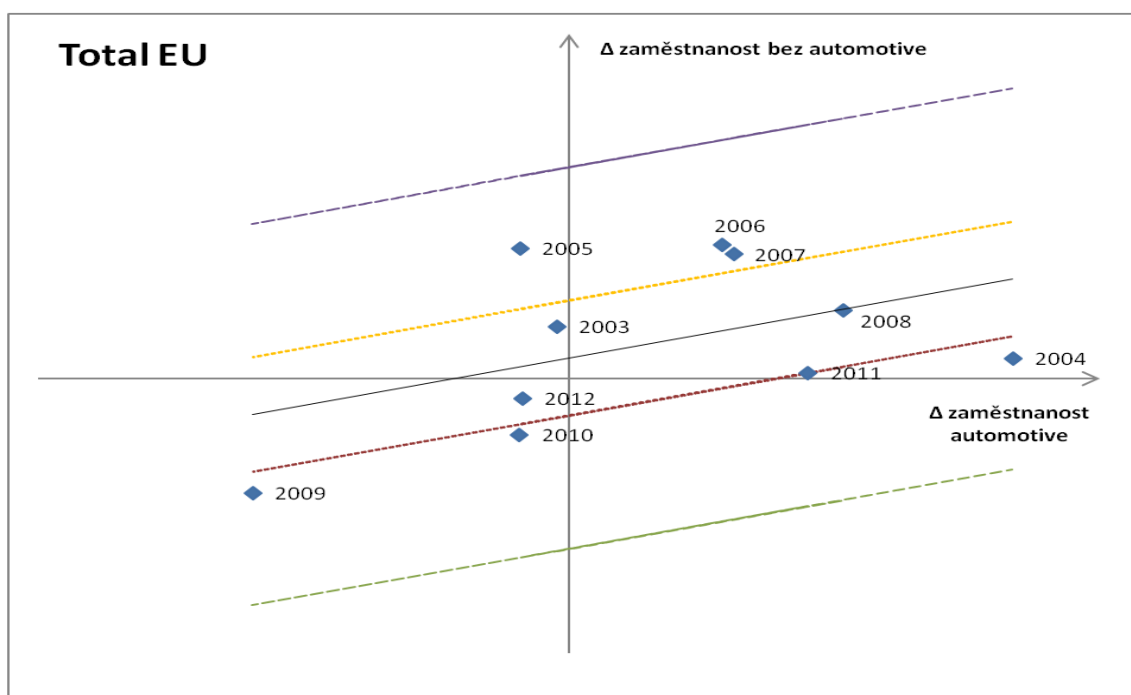
Tabulka 4.10 Výsledek lineární regrese změn zaměstnanosti mimo automobilový průmysl na změnách zaměstnanosti v automobilovém průmyslu v EU

Hodnota F	0,02026
Významnost F	0,89033
Hladina významnosti α	0,05
Koeficient β_1	-0,01481
Korelační koeficient	0,05026

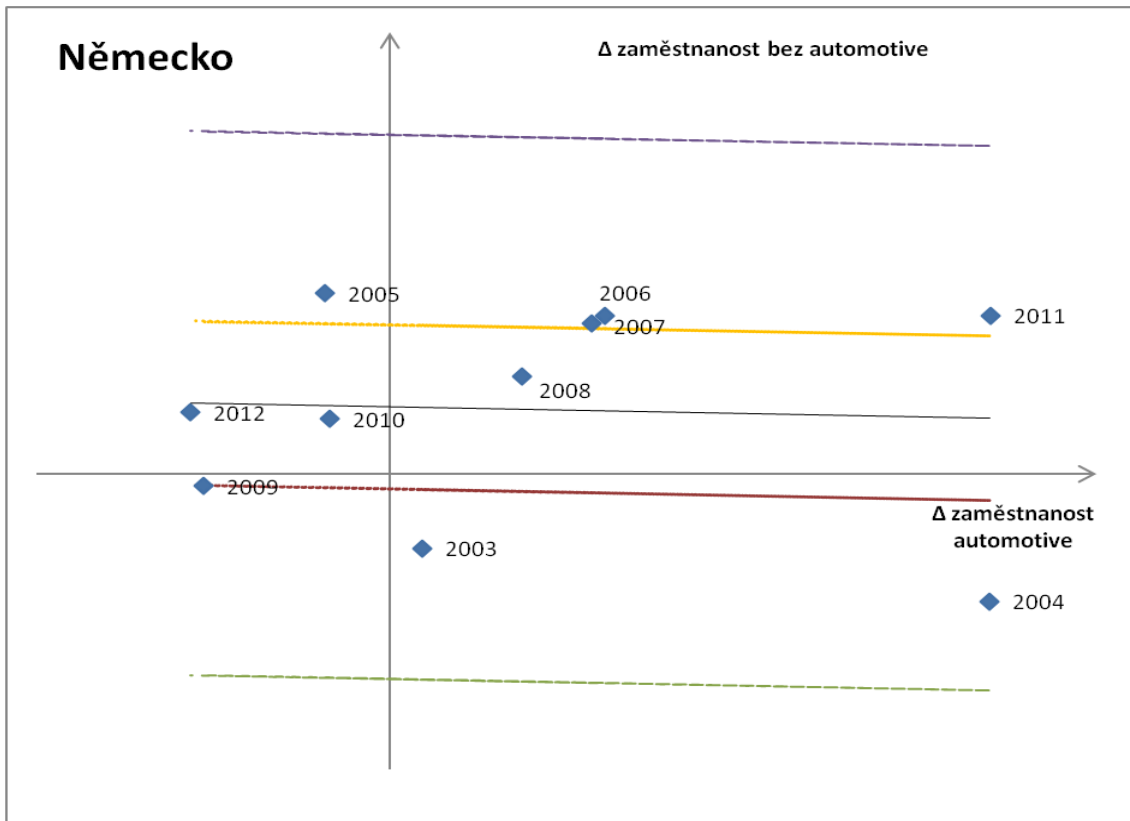
Tabulka 4.11 Výsledek lineární regrese změn zaměstnanosti mimo automobilový průmysl na změnách zaměstnanosti v automobilovém průmyslu v Německu

Hypotézu H_0 nezamítáme na hladině významnosti $\alpha = 5\%$ pro země EU celkem a také pro Německo, model je statisticky nevýznamný. Závislost změn zaměstnanosti v automobilovém průmyslu v EU a také v Německu na změnách zaměstnanosti mimo automobilový průmysl se nám tedy nepodařilo prokázat. Jak jsme však zmínili výše, může být výsledek ovlivněn malým počtem pozorování.

Graf lineární regrese pro země EU celkem resp. pro Německo je znázorněn na obrázku 4.4 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v EU resp. 4.5 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v Německu.



Obrázek 4.4 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v EU



Obrázek 4.5 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v Německu

4.5 Intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování

Pro modely uvedené v předchozím odstavci jsme se rozhodli zkonstruovat intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování, tzv. pásy spolehlivosti. (Reif, 2004)

Nestranné odhady rozptylů hodnot $b_0 + b_1x_0$ a Y jsou rovny:

$$(1) \quad \widehat{D}(b_0 + b_1x_0) = s_R^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{n\sigma_n^2(x)} \right]$$

$$(2) \quad \widehat{D}(Y) = \widehat{D}(b_0 + b_1x_0) + s_R^2$$

kde σ^2 je rozptyl statistického souboru a s_R^2 je reziduální rozptyl.

Odhady směrodatných odchylek jsou:

$$(3) \quad SE(b_0 + b_1x_0) = \sqrt{\widehat{D}(b_0 + b_1x_0)}$$

$$(4) \quad SE(Y) = \sqrt{\widehat{D}(Y)}$$

(Reif, 2004)

Tyto intervaly jsou zobrazeny na obrázku 4.4 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v EU resp. 4.5 Lineární regrese změn zaměstnanosti v automotive na změnách zaměstnanosti mimo automotive v Německu. Hodnoty změn zaměstnanosti mimo automobilový průmysl se tedy s pravděpodobností 95%

nacházejí uvnitř užšího pásu, budoucí pozorování se budou se stejnou pravděpodobností nacházet v širším pásu (predikce).

Intervaly spolehlivosti pro regresní hodnoty $SE(b_0 + b_1x_0)$ a pro budoucí pozorování $SE(Y)$, tedy nestranné odhady rozptylů hodnot, vypočtené podle výše uvedených vzorců (1), (2), (3) a (4) jsou uvedeny v tabulce 4.12 Intervaly spolehlivosti pro regresní hodnoty $SE(b_0 + b_1x_0)$ a pro budoucí pozorování $SE(Y)$ pro EU a 4.13 Intervaly spolehlivosti pro regresní hodnoty $SE(b_0 + b_1x_0)$ a pro budoucí pozorování $SE(Y)$ pro Německo.

intervaly spolehlivosti					
	<i>dolní</i> $SE(Y)$	<i>horní</i> $SE(Y)$		<i>dolní</i> $SE(b_0 + b_1x_0)$	<i>horní</i> $SE(b_0 + b_1x_0)$
-	0,02513	0,03036	-	0,00575	0,01098
-	0,01334	0,04215	-	0,00604	0,02277
-	0,02610	0,02939	-	0,00672	0,01001
-	0,02087	0,03462	-	0,00149	0,01524
-	0,02056	0,03493	-	0,00118	0,01555
-	0,01773	0,03776	-	0,00165	0,01838
-	0,03302	0,02247	-	0,01364	0,00309
-	0,02611	0,02937	-	0,00674	0,01000
-	0,01864	0,03685	-	0,00074	0,01747
-	0,02602	0,02947	-	0,00664	0,01009

Tabulka 4.12 Intervaly spolehlivosti pro regresní hodnoty $SE(b_0 + b_1x_0)$ a pro budoucí pozorování $SE(Y)$ pro EU

intervaly spolehlivosti					
	<i>dolní</i> $SE(Y)$	<i>horní</i> $SE(Y)$		<i>dolní</i> $SE(b_0 + b_1x_0)$	<i>horní</i> $SE(b_0 + b_1x_0)$
-	0,02786	0,04601	-	0,00207	0,02021
-	0,02929	0,04458	-	0,00349	0,01878
-	0,02762	0,04625	-	0,00182	0,02045
-	0,02832	0,04555	-	0,00253	0,01975
-	0,02829	0,04558	-	0,00249	0,01978
-	0,02811	0,04576	-	0,00231	0,01996
-	0,02731	0,04656	-	0,00151	0,02076
-	0,02763	0,04624	-	0,00183	0,02044
-	0,02929	0,04458	-	0,00349	0,01878
-	0,02728	0,04659	-	0,00148	0,02079

Tabulka 4.13 Intervaly spolehlivosti pro regresní hodnoty $SE(b_0 + b_1x_0)$ a pro budoucí pozorování $SE(Y)$ pro Německo

4.6 Lineární regrese produkce osobních a nákladních automobilů v EU a ověřování předpokladů regrese

V rámci této kapitoly jsme se ještě rozhodli ověřit předpoklady lineární regrese.

Předpoklady budeme testovat na čtvrtletních hodnotách indexu produkce osobních automobilů v EU jako vysvětlující proměnnou (x_i) a indexu produkce nákladních automobilů v

EU jako proměnnou vysvětlovanou (y_i), kde index 100 je roven hodnotám produkce v roce 2010. Čtvrtletní data pocházejí z let 2002 až 2013.

Mějme opět následující model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Pro náhodné odchyly tohoto regresního modelu předpokládáme následující podmínky:

1. Střední hodnota náhodné odchyly $E(\varepsilon_i)$ je rovna 0
2. Náhodná odchylnka má konstantní rozptyl $D(\varepsilon_i)$ pro každé pozorování
3. Veličiny ε_i jsou nezávislé. (Reif, 2004)

Porušení každé z těchto podmínek značí jisté problémy pro odhady parametrů nebo intervaly spolehlivosti. (Statistica, 2015)

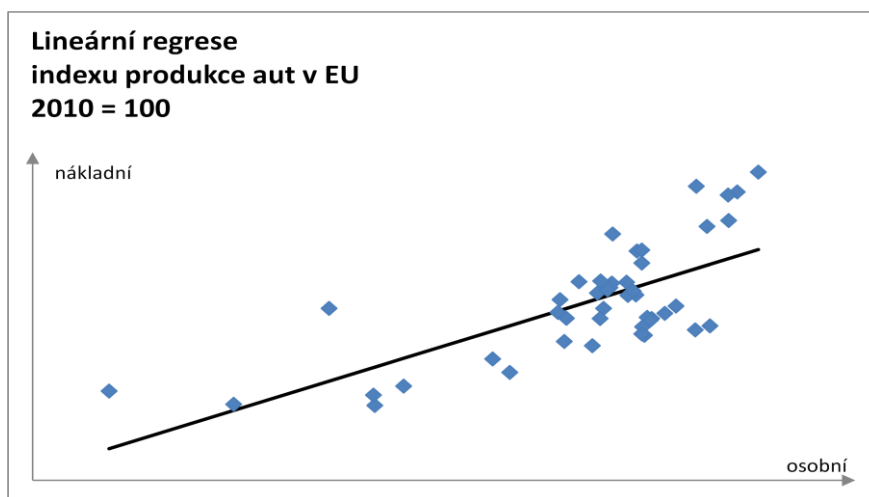
Pro uvedený regresní model jsme stanovili hypotézu:

$$H_0 : \beta_1 = 0,$$

kterou zamítáme na hladině významnosti $\alpha = 5\%$. Nezávislost mezi produkcí osobních a nákladních automobilů v EU v letech 2002-2013 se nám nepodařilo prokázat. Výsledek regresní analýzy a grafické zobrazení je uvedeno v tabulce 4.14 Výsledek lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013 a na obrázku 4.6 Lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013.

Hodnota F	43,70517
Významnost F	$4,72059 \cdot 10^{-8}$
Hladina významnosti α	0,05
Koeficient β_1	1,79607
Korelační koeficient	0,70998

Tabulka 4.14 Výsledek lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013



Obrázek 4.6 Lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013

Nyní se blíže podíváme na testování výše uvedených předpokladů lineární regrese.

4.6.1 Předpoklad nulovosti střední hodnoty náhodných odchylek

1. předpoklad nulovosti střední hodnoty náhodných odchylek ε_i jsme testovali jednovýběrovým testem o průměru při neznámém rozptylu, t-testem. Formulovali jsme následující nulovou a alternativní hypotézu:

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

Testovací kritérium

$$T = \frac{\bar{x} - \mu_0}{s} \sqrt{n}$$

v absolutní hodnotě jsme porovnali s příslušnými kvantily Studentova t-rozdělení se stupni volnosti $(n - 1)$. (Šedivá & Ťoupal, 2012) Výsledky testu jsou uvedeny v tabulce 4.15 Výsledek t-testu o průměru při neznámém rozptylu.

s	n	\bar{x}	μ_0	α	$t_{1-\alpha/2}(n-1)$	T
17,08	45	0,0	0	5%	2,02	0,00

Tabulka 4.15 Výsledek t-testu o průměru při neznámém rozptylu

Na základě výsledků nezamítáme nulovou hypotézu $H_0 : \mu = 0$ na hladině významnosti $\alpha = 5\%$ a lze tedy říci, že náhodné odchylky ε_i s pravděpodobností 95% pochází ze stejného rozdělení se střední hodnotou rovnou nule.

4.6.2 Předpoklad konstantního rozptylu náhodných odchylek

Pro testování 2. předpokladu, tedy homoskedasticitu resp. heteroskedasticitu modelu, jsme zvolili Goldfeld-Quandtův test. (Reif, 2004)

Testujeme nulovou a alternativní hypotézu:

$$H_0 : D(\varepsilon_i) = \sigma^2$$

$$H_1 : D(\varepsilon_i) \neq \sigma^2$$

Testovací statistika má tvar:

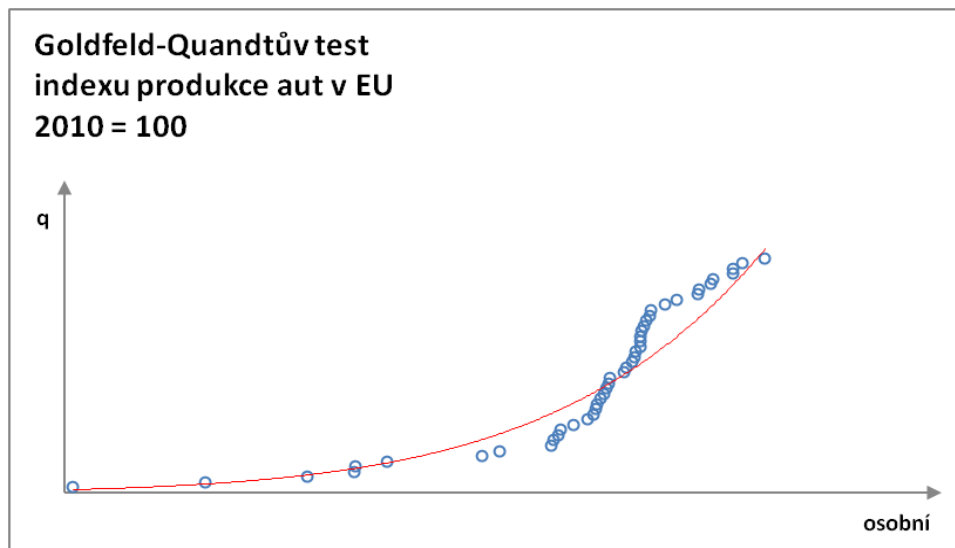
$$F = \frac{s_{R1}^2}{s_{R2}^2}$$

kde s_{R1}^2 , s_{R2}^2 jsou příslušné reziduální rozptyly.

Na základě výsledků testu uvedených v tabulce 4.16 Výsledek Goldfeld-Quandtova testu o heteroskedasticitě modelu nezamítáme hypotézu o homoskedasticitě na hladině významnosti $\alpha = 5\%$, a tvrdíme, že heteroskedasticita modelu je statisticky nevýznamná. Rozptyl náhodných odchylek modelu je konstantní.

n	q	α	s^2_{R1}	s^2_{R2}	F	$F_{1-\alpha/2}(q-2, q-2)$
45	15	5%	2 467,30	5 295,23	2,15	2,16

Tabulka 4.16 Výsledek Goldfeld-Quandtova testu o heteroskedasticitě modelu



Obrázek 4.7 Goldfeld-Quandtův test indexu produkce osobních a nákladních aut v EU v letech 2002-2013

4.6.3 Předpoklad nezávislosti náhodných odchylek

Při splnění 3. předpokladu, tzn. nezávislosti náhodných odchylek, by korelační koeficient odchylek ε_i byl roven nule. Nejjednodušším modelem korelace odchylek je tzv. autokorelace 1. řádu, jejíž nulovost ($\rho = 0$) jsme prověřili Durbin-Watsonovým testem. (Reif, 2004)

Hodnoty porovnááme s testovací statistikou:

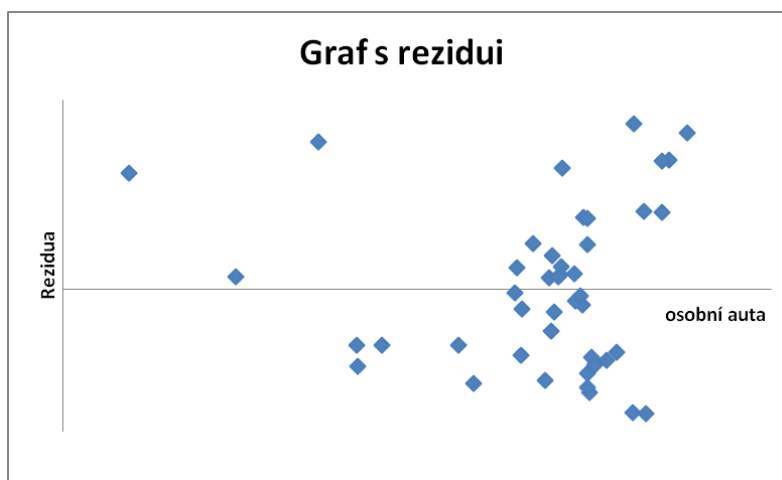
$$D = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Na základě porovnání hodnot testovací statistiky D (Novák, 2005) s kritickou hodnotou zamítáme hypotézu $H_0: \rho = 0$ na hladině významnosti $\alpha = 5\%$, D patří do intervalu $< 0; d_d >$, jedná se tedy o pozitivní autokorelaci. Výsledek testu viz tabulka 4.17 Výsledek Durbin-Watsonova testu nulovosti autokorelace 1. řádu.

n	α	$\sum_{i=2}^n (e_i - e_{i-1})^2$	$\sum_{i=1}^n e_i^2$	D	d_d	d_h
45	5%	18 110,62	12 833,28	1,41	1,48	1,57

Tabulka 4.17 Výsledek Durbin-Watsonova testu nulovosti autokorelace 1. řádu

Nezávislost reziduí ε_i můžeme také ilustrovat na obrázku 4.8 Zobrazení reziduí lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013.



Obrázek 4.8 Zobrazení reziduí lineární regrese indexu produkce osobních a nákladních aut v EU v letech 2002-2013

5 Závěr

Tato práce se zaměřila na statistické zpracování dat z oblasti automotive. Pro tuto analýzu byla shromážděna data z několika ekonomických oblastí jako je HDP, počet obyvatel, HDP/obyvatele, zaměstnanost v zemích EU, zaměstnanost v automobilovém průmyslu, produkce osobních a nákladních automobilů a index produkce osobních a nákladních automobilů. Datový soubor byl získán z databáze evropského statistického úřadu Eurostat a data byla vyhledávána ze všech 28 členských států Evropské unie mezi roky 2000 - 2014.

V kapitole č. 3 byly nejprve určeny průměry a mediány jako stanovení základních statistických charakteristik v jednotlivých zemích EU a pro sledované roční ekonomické ukazatele. V této kapitole jsme při základní analýze zároveň formulovali domněnku závislosti HDP/obyvatele a produkce aut/obyvatele v letech 2006-2012 v České republice. Zajímala nás také závislost HDP/obyvatele na produkci aut v EU ve vybraných obdobích, která mohla odpovídat pozitivnímu ekonomickému vývoji v průběhu let 2003-2012. Na vývoji produkce osobních a nákladních automobilů v jednotlivých čtvrtletích let 2003-2014 jsme sledovali propad produkce, který mohl být odezvou na ekonomickou krizi.

Ve čtvrté kapitole jsme pomocí lineární regrese testovali závislost výše reálného HDP/obyvatele v České republice na produkci v automobilovém průmyslu. Vzhledem k výsledkům analýzy se jednalo o silnou kladnou závislost, což nás přivedlo k domněnce, že se může jednat o klamnou korelaci. V následujícím odstavci jsme testem nezávislosti potvrdili, že výše reálného HDP/obyvatele v ČR závisela na produkci v automobilovém průmyslu s vyloučením vlivu času. Při testování jsme zjistili, že v určitém bodě se křivka závislosti mění a že bod odpovídá počátku ekonomické krize v roce 2008. Data jsme tedy po rozdělení na dobu před ekonomickou krizí a v krizi znovu prověřili. Na základě výsledků se nám v letech 2000-2007, tedy v době před ekonomickou krizí, nepodařilo prokázat, že data jsou v čase závislá a tedy klamně korelovaná, ale pro roky 2008 až 2009, tedy dobu ekonomické krize, jsme konstatovali, že se jednalo o klamnou korelaci. V dalším odstavci této kapitoly jsme ověřovali hypotézu korelace změn zaměstnanosti v automobilovém průmyslu a mimo něj pro země EU a pro Německo a také jsme zkonstruovali intervaly spolehlivosti pro regresní hodnoty a pro budoucí pozorování. Závislost změn zaměstnanosti v automobilovém průmyslu a mimo něj se nám na základě výsledků nepodařilo prokázat. V posledním odstavci čtvrté kapitoly jsme na čtvrtletních hodnotách indexu produkce osobních automobilů v EU a indexu produkce nákladních automobilů v EU z let 2002 až 2013 po provedení testu závislosti dat, ověřovali předpoklady lineární regrese. Nezávislost mezi produkcí osobních a nákladních automobilů v EU v letech 2002-2013 se nám nepodařilo prokázat, nicméně po provedení Durbin-Watsonova testu jsme prokázali pozitivní autokorelaci reziduí a tím porušení 3. předpokladu lineární regrese.

Prokázat, že automotive je skutečně motorem ekonomiky v ČR, Německu nebo v jiných státech EU a že jeho ekonomický vývoj má podstatný vliv na fungování celého tržního systému se nám sice na základě výsledků této práce nepodařilo, nicméně z důvodu

omezeného množství dat, které jsme měli o této oblasti k dispozici, nelze tvrzení ani zamítnout.

6 Citovaná literatura

Český statistický úřad. (Duben 2015). <https://www.czso.cz/> .

Novák, P. (2005). Kritické hodnoty pro Durbin-Watsonův test (5%-ní hladina významnosti).

Pavlínek, P., & Ženka, J. (2010). The 2008–2009 automotive industry crisis and regional unemployment in Central Europe. *Cambridge Journal of Regions* .

Reif, J. (2004). *Metody matematické statistiky*. Plzeň: Západočeská univerzita.

Statistica. (Květen 2015). *Diagnostika regrese pomocí grafu 7krát jinak*. Načteno z www.stasoft.cz

Statistický úřad Evropské unie. (Duben 2015). <http://ec.europa.eu/eurostat> .

Sůra, J. (2011). *Motor ekonomiky chladne. O nová auta přestává být v Evropě zájem*. Získáno 06 2015, z iDnes.cz/Ekonomika: <http://ekonomika.idnes.cz/motor-ekonomiky-chladne-o-nova-auta-prestava-byt-v-evrope-zajem-ps4/>

Šedivá, B., & Toupal, T. (2012). *Přednáška k předmětu Výpočtová statistika* .

7 Obsah přiloženého CD

1. *Bakalářská práce* ve formátu pdf
2. Adresář *Zpracování* obsahující soubory ve formátu xls, ve kterých bylo provedeno statistické zpracování dat, jejich grafické zobrazení a tabulky s výsledky
3. Adresář *Zdrojová data* obsahující shromážděná data ve formátech .xls a .pdf